Community Capability Model Interest Group Meeting

RDA 3nd Plenary Meeting
Thursday 27th March, 2014
Dublin, Ireland

Manjula Patel
UKOLN Informatics
University of Bath, UK

Acknowledgements

This work is funded by Microsoft Research Connections.

UKOLN Informatics receives additional support from the University of Bath where it is based.

Contacts:

Liz Lyon: elyon@pitt.edu

Kenji Takeda: kenjitak@microsoft.com

Manjula Patel: m.patel@ukoln.ac.uk

Further Information:

http://communitymodel.sharepoint.com/

Current Interest Group Goals

- To demonstrate application of the CCMF Profile tool across a range of contexts and communities: disciplines, organisations, groups etc.
- To collect completed profiles from researchers in a diverse range of disciplines and sub-disciplines
- To investigate opportunities to customise the Profile template for particular disciplinary domains

Meeting Agenda

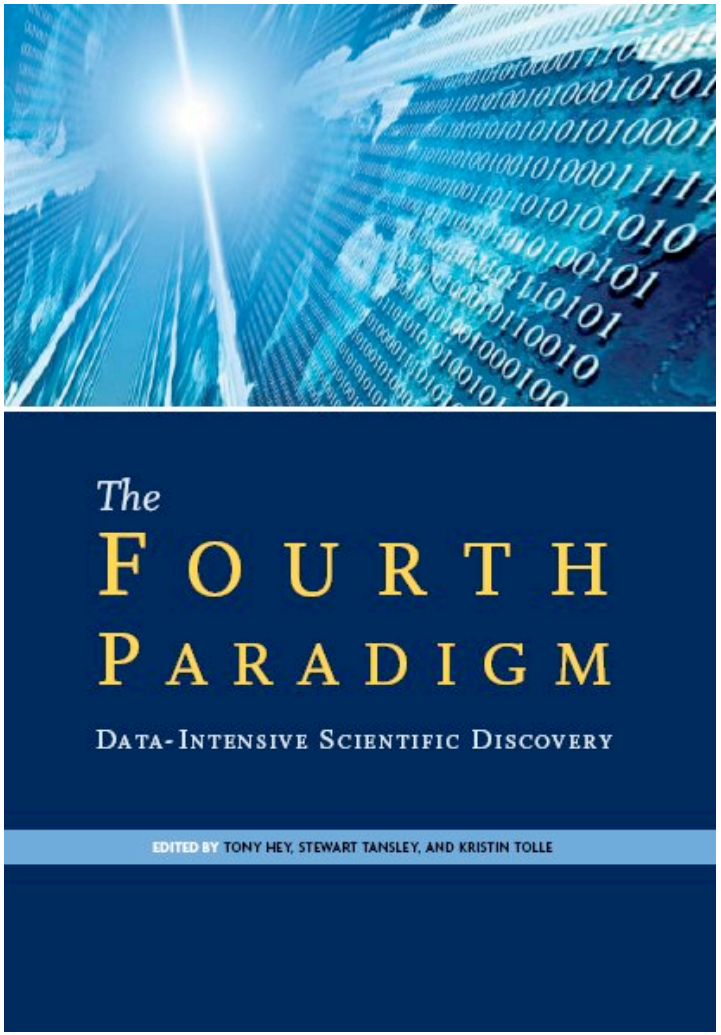
- Review of CCMF Profile tool (for newcomers)
- Developments since 2nd RDA Plenary Meeting
  - Use case: Agronomy
  - Use case: DataONE
- Issues
- Discussion and Next steps
- AOB

Context

- Experimental Science
  - Observational description of natural phenomena
- Theoretical Science
  - Use of models and equations

    e.g. Newton's Laws
- Computational Science
  - Digital simulation of complex phenomena
- Data-Intensive Science
  - Unify experiment, theory and simulation

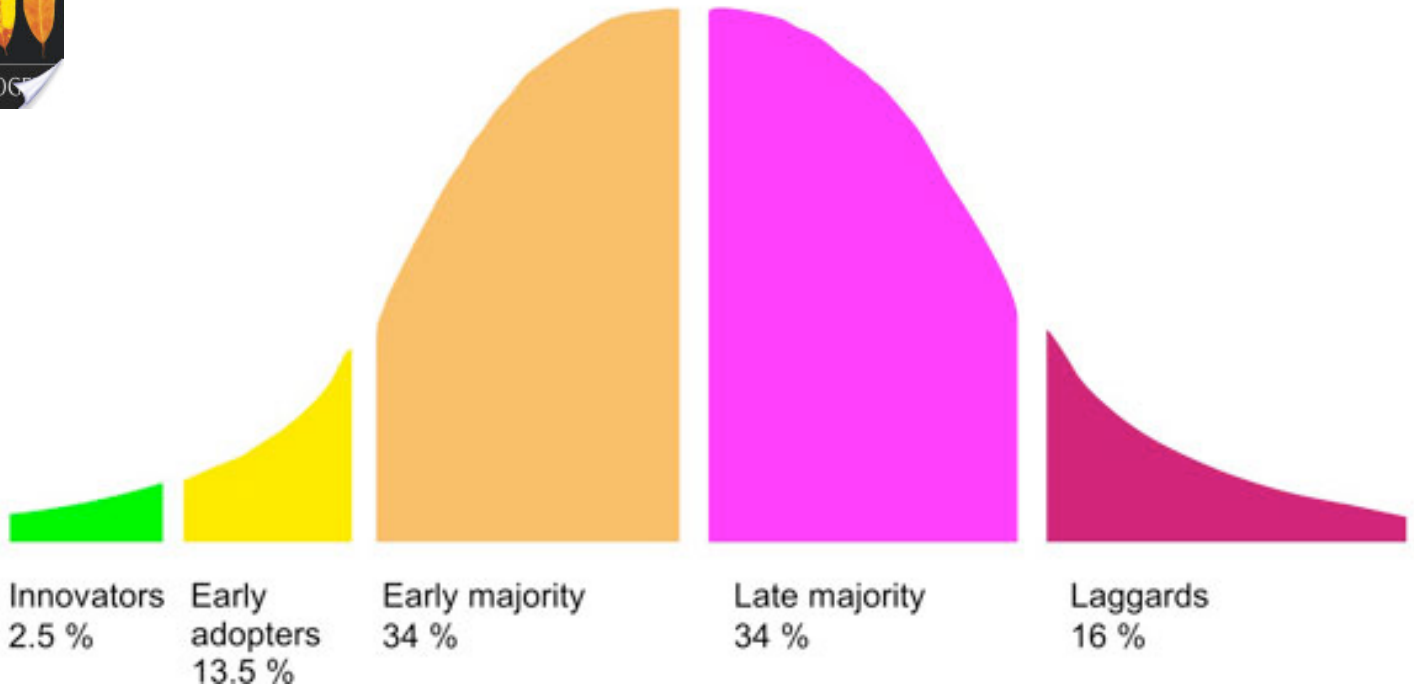
    - Jim Gray
- Data-Intensive Research
  - Intensive data collection and processing
  - Big Data
  - Aggregation of diverse datasets



research.microsoft.com/en-us/collaboration/fourthparadigm/

Diffusion of Data-Intensive Research

Everett M. Rogers, *Diffusion of Innovations* , Fifth Edition 2003, Free Press, New York
http://books.google.co.uk/books/about/Diffusion_of_Innovations_5th_Edition.html?id=9U1K5LjUOwEC



www.neb-one.gc.ca/clf-nsi/rnrgynfmtn/nrgyrprt/nrgdmnd/tttdbhvrshpngnrgs2009/mg/f04-eng.jpg

Motivations for DIR

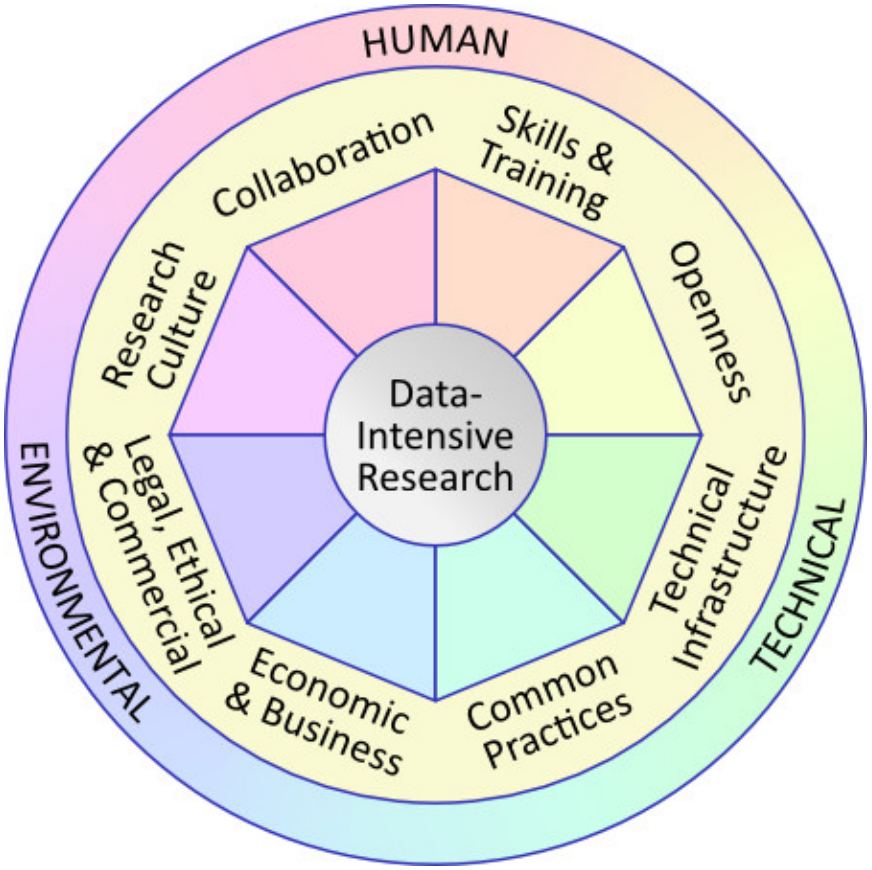
- Funding Bodies (e.g. NSF, European Union, UK Research Councils, Trusts, Learned Societies, Companies, Foundations)
  - Derive maximum research, economic and social benefits from investments
  - Improve the quality and efficiency of research (robust and reproducible)
  - Increase knowledge transfer within discipline; across disciplines; between sectors
  - Build sub-disciplinary, disciplinary and inter-disciplinary communities
  - Develop added-value services based on corpora of research data
- Institutions (e.g. HEIs, Facilities (e.g. CERN, STFC, EMBL))
  - Improve the quality and efficiency of research (robust and reproducible)
  - Increase ability to attract research funds
  - Build institutional and cross-institutional communities
  - Develop added-value services based on corpora of research data
  - Include data citation into research evaluation systems e.g. UK's REF
- Researchers (Principal Investigators)
  - Opportunities for new and innovative research
  - Improve the quality of research (robust and reproducible)
  - Improve citations and reputation
  - Career advancement
  - Add data citation into research evaluation systems e.g. UK's REF

Data-Intensive Research Lifecycle

Areas that need particular attention

- Legal, ethical and commercial issues
  - IPR, privacy, sensitivity, licensing
- Gaining informed consent for reuse and repurposing
- Appraisal and quality control
  - Collection and acquisition policies, peer review
- Trustworthiness
  - Metadata, documentation, context, provenance, transparency
- Scale and complexity of data
  - Workflows, methodologies, software, OAIS Representation Information
- Publication and sharing
  - Release policy, controlled access (embargoes), indexing, interoperability (syntax and semantics), cross-searching, federation
- Citation, attribution and accreditation in scholarly communications
  - granularity, versioning, persistent identifiers

The CCMF



communitymodel.sharepoint.com

- The Community Capability Model Framework (CCMF)
  - Profiling current readiness or capability of a community for DIR
  - Indicating priority areas for change and investment
  - Developing roadmaps for achieving a target state of readiness

CCMF White Paper, April 2012

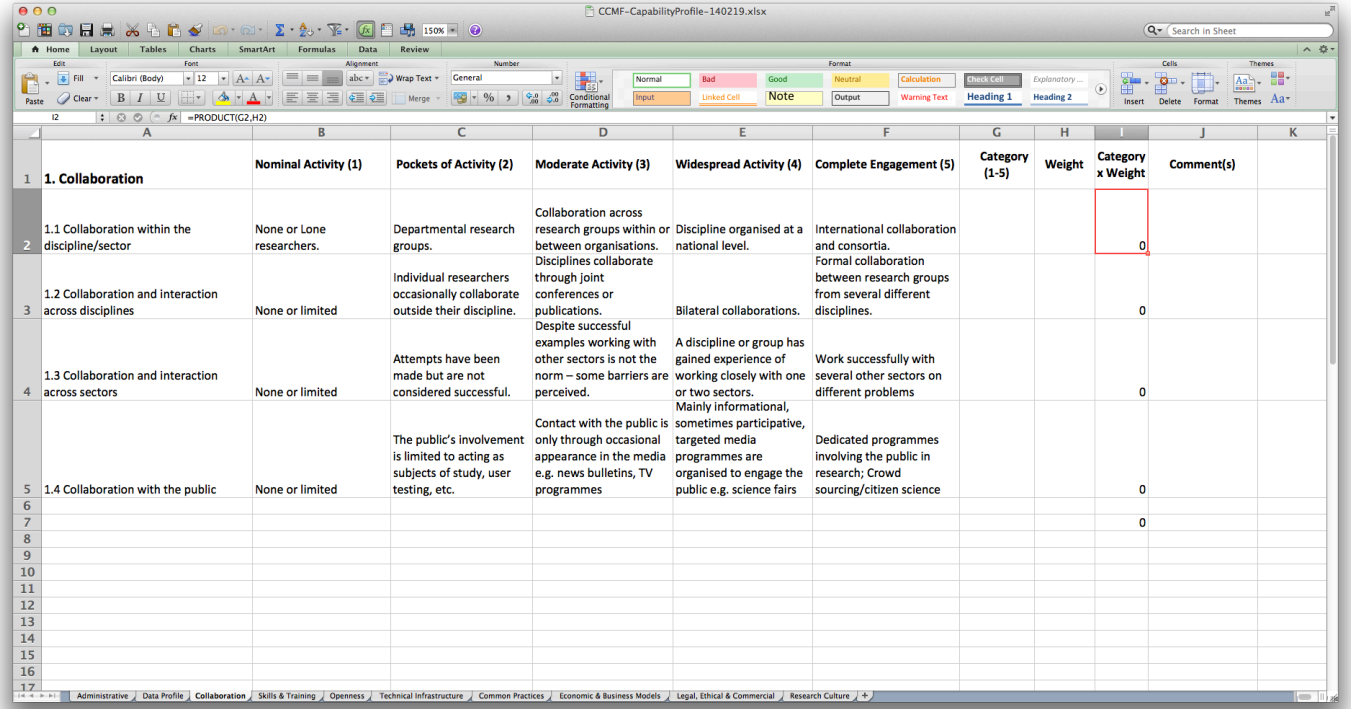
- Developed through consultation: case studies and workshops
- Primarily a tool for self-assessment and longitudinal studies
- Categorised into Environmental, Human and Technical elements with eight factors:

Openness	Legal, Ethical & Commercial
Collaboration	Economic & Business
Skills & Training	Common Practices
Research Culture	Technical Infrastructure

- Each factor has *characteristics* associated with it

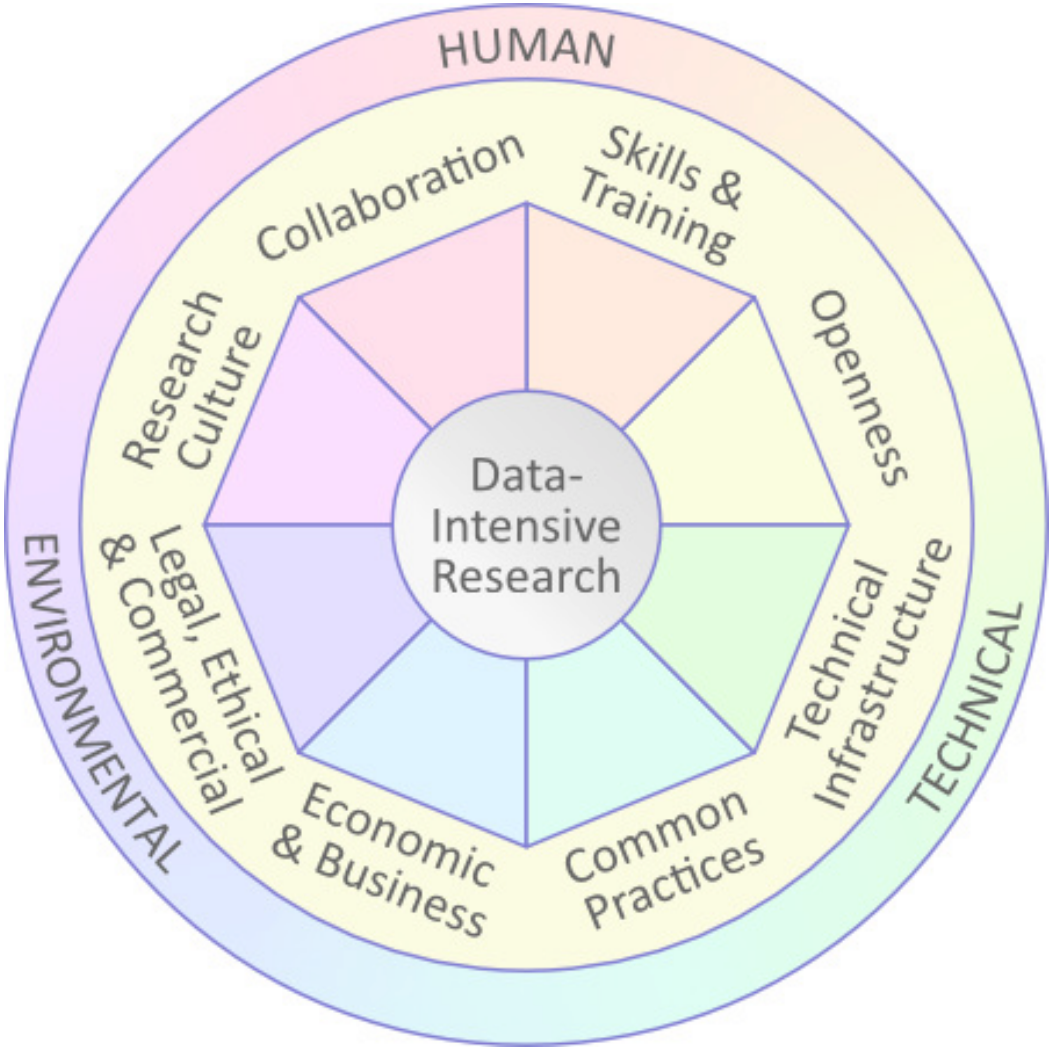
CCMF Profile Tool

- Implemented as an MS Excel spread sheet
- Separate worksheets for each of the eight CCMF factors
- A scorecard tool (5 levels or *dimensions* for each *characteristic* within each *factor*)
- Download CCMF Profile Template from:
  https://communitymodel.sharepoint.com/



	A	B	C	D	E	F	G	H	I	J	K
1	**1. Collaboration**	**Nominal Activity (1)**	**Pockets of Activity (2)**	**Moderate Activity (3)**	**Widespread Activity (4)**	**Complete Engagement (5)**	**Category (1-5)**	**Weight**	**Category x Weight**	**Comment(s)**	
2	1.1 Collaboration within the discipline/sector	None or Lone researchers.	Departmental research groups.	Collaboration across research groups within or between organisations.	Discipline organised at a national level.	International collaboration and consortia.			0		
3	1.2 Collaboration and interaction across disciplines	None or limited	Individual researchers occasionally collaborate outside their discipline.	Disciplines collaborate through joint conferences or publications.	Bilateral collaborations.	Formal collaboration between research groups from several different disciplines.			0		
4	1.3 Collaboration and interaction across sectors	None or limited	Attempts have been made but are not considered successful.	Despite successful examples working with other sectors is not the norm – some barriers are perceived.	A discipline or group has gained experience of working closely with one or two sectors.	Work successfully with several other sectors on different problems			0		
5	1.4 Collaboration with the public	None or limited	The public's involvement is limited to acting as subjects of study, user testing, etc.	Contact with the public is only through occasional appearance in the media e.g. news bulletins, TV programmes	Mainly informational, sometimes participative, targeted media programmes are organised to engage the public e.g. science fairs	Dedicated programmes involving the public in research; Crowd sourcing/citizen science			0		
6											
7									0		
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											

CCMF Profile Tool Worksheets



Recent Developments

- IDCC 2014 workshop - Delegates wanted to
  - Change language to be more (sub)discipline specific
  - Change examples to be more relevant to their own domain
- Use case: Data Observation Network for Earth (DataONE)
  - An umbrella organisation covering all Environment Science
  - One Profile completed collectively by SMT on behalf of whole organisation
  - Need IRB approval to disseminate Profile to partners
- Use case: Agronomy (Purdue University)
  - Agronomy specific Profile template
  - Three Profiles completed by three individual agronomists
  - Need IRB approval for widespread dissemination and collection of completed Profiles

Agenda

- Developments since 2nd RDA Plenary Meeting
  - Use case: Agronomy
  - Use case: DataONE
- Review of CCMF Profile tool
- Issues
  - Who would like their completed Profiles to be published anonymously?
  - Who needs IRB approval before they are prepared to fill in the profile?
- Discussion and Next Steps
  - Domain “champions” to undertake localisation of Profile template
  - Collect lots of completed profiles for analysis and comparison
  - Effective visualisation of results and comparisons
  - Platform for community engagement
- AOB